DOCUMENT RESUME

ED 276 746 TM 860 696

AUTHOR Katzenmeyer, Conrad G.; Haertel, Geneva
TITLE Analyzing the JDRP as an Evaluation Process.
PUB DATE Apr 86
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Quality; Elementary Secondary Education; *Evaluation Criteria; *Evaluation Methods; *Program Evaluation; *Standards
IDENTIFIERS Effective Schools Research; *Joint Dissemination Review Panel; National Science Teachers Association; *Standards for Evaluation Educ Prog Proj Materials

ABSTRACT

The Joint Dissemination Review Panel (JDRP), part of the U.S. Department of Education, is one of the most important federal mechanisms for recognizing quality in education. Although it has reviewed close to 700 projects in its 15-year history, the JDRP is still not widely known in the field of evaluation. This paper analyzes the JDRP process and its evaluation practices and contrasts it with evaluation criteria of the National Science Teachers Association and with effective schools research. This paper further analyzes the JDRP through the use of the Standards for Evaluations of Educational Programs, Projects and Materials, developed by the Joint Committee on Standards for Educational Evaluation. The standards are grouped in four areas: usability, feasibility, propriety, and accuracy. Results highlight those elements of an evaluation to which the JDRP attends and how this has influenced the composition of projects it reviews and approves. (Author/JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED276746

Analyzing the JDRP as an Evaluation Process

by

Conrad G. Katzenmeyer
Office of Educational Research and Improvement
U.S. Department of Education

and

Geneva Haertel
Stanford, California

Abstract

Despite having reviewed close to 700 projects over a 15 year history, the U.S. Department of Education's Joint Dissemination Review Panel is still not widely known in the field of evalution. There have also been few studies analyzing the JDRP process. The purpose of this study was to contrast the JDRP with several other recognition processes, and to analyze the JDRP through the use of the Standards for Evaluations of Educational Programs, Projects and Materials. Results highlight those elements of an evaluation to which the JDRP attends, and how this has influenced the composition of projects it reviews and approves.

Paper presented at the American Educational Research Association meetings, San Francisco, California, April, 1986

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Conrad G. Katzenmeyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 860 696

Analyzing the JDRP as an Evaluation Process

by

Conrad G. Katzenmeyer
Office of Educational Research and Improvement
U.S. Department of Education

and

Geneva Haertel
Stanford, California

Recognition of educational quality, whether in students, programs, or schools, has always been of interest to the field of education, and particularly to the U.S. Department of Education. However, it has received renewed attention with the current administration's concern for excellence. There now exist a number of programs charged with the responsibility of identifying educational quality.

Certainly one of the most important federal mechanisms for recognizing quality in education over the past decade has been the Joint Dissemination Review Panel. Given the large number of projects that have been reviewed by the JDRP, as well as its role with NDN, it is our belief that the JDRP represents one of the most significant examples of the use of evaluation in education. It is therefore worthwhile to analyze what this process has been and what it tells us about the practice of educational evaluation. We have chosen to do this by contrasting the JDRP process with two other approaches that share some of the purposes of the JDRP, and to match the JDRP to a set of standards for program evaluation that are now generally accepted in the field of educational evaluation.

Alternate Approaches for Recognizing Quality

Every recognition system is bound to be influenced by the nature of the candidate projects it reviews. For the JDRP, a major source of projects until recently was Title III(Title IVc later) of the Elementary and Secondary Education Act. These projects had several distinguishing characteristics. They were generally large-scale and well funded, both for development and evaluation. They also placed major emphasis on the development of materials and processes; thus, they generally generated a tangible product. Finally, the emphasis was on innovation and change. The prevailing model was the social science paradigm for social programs, with the accompanying social science approach to verification.

The approach to evaluation employed in the JDRP process is the one now labeled "program evaluation" in the literature. In order to meet the requirements for an evaluation of this type, the program must be a narrow, clearly specified intervention. It must be implemented in a standard, verifiable manner. And it must have objective, measurable outcomes. The last of these is critical for the JDRP; outcomes are the parts of the program that can be addressed most easily and unambiguously by experimental social science methodology. This doesn't necessarily imply an emphasis on quantitative methods and standardized achievement tests, but for the most part, this is what happened with the JDRP.

However, there are other criteria and techniques that have been used for recognizing quality. An interesting contrast is the approach that the National Science Teachers Association used in its Focus on Excellence. The purpose was to identify exemplary school science programs. Unlike projects in some curricular areas, science projects tend to emphasize process variables, and seldom rely on standardized achievement tests as outcome measures. The projects were nominated and initially reviewed with a process similar to JDRP's, but site visits were used in the final selection. Emphasis was placed on content innovation, with outcomes determined by content specialists' judgments rather than tests and quasi-expermimental designs. This approach is quite similar to what NSF used with its curriculum development projects of the 1960s and early 1970s. Despite the differences in procedures employed, however, this is also a program evaluation and the purpose is the same as for the JDRP.

How do these approaches compare? The 50 projects selected for Focus on Excellence in 1982 were retrospectively reviewed, using the JDRP criteria(Katzenmeyer, 1985.) The results indicate that if the 50 projects had been submitted to the JDRP, two were likely to be passed(one had in fact been approved), eight were doubtful, and 40 were very unlikely to receive JDRP approval. In other words, of 50 nationally acclaimed science programs, between 80 and 96% would fail JDRP.

But what about this comparison from the opposite perspective? Would projects that pass JDRP have made it through Focus on Excellence or the NSF review? No direct analysis is possible because so few science projects have been passed by JDRP, and several of these were funded by NSF. But our experience in evaluating a large number of Title III projects and in participating in the NSF review process lead us to believe that very few Title III projects would have been considered worthy by NSF. NSF's emphasis on high level science content, with great attention to accuracy and comprehensiveness, was simply beyond the capabilities of most Title III projects.

The difference between the NSF and JDRP approaches is a matter of values. NSF valued state-of-the-art content; whether students learned science became the responsibility of the teacher and local school in how they used the content. JDRP values student outcomes; whether this is on state-of-the-art content is the

responsibility of the teacher and local school in selecting the program. While this somewhat overstates the value conflict, it is clear that there have been strikingly few examples of projects that contain both superior content and superior demonstrated outcomes in science education.

Another contrast to the JDRP approach is the identification process used in the effective schools literature. The concept of effective schools began with the observation that students in a few schools showed levels of academic achievement, as indicated by standardized tests scores, in excess of what was being achieved in the majority of schools serving the same kinds of students.

Schools with achievement levels substantially higher than statistical predictions, based on student characteristics, were identified as effective schools. Researchers then studied these schools on a case by case basis to find the keys to their success. Effective schools were found to be those with a strong schoolwide press toward academic learning, principals who were strong instructional leaders, clearly articulated and consistently enforced disciplinary and attendance policies, and reliance on teacher-centered didactic instruction.

Clearly, the test of the JDRP in identifying effective programs based on treatment-control group comparisons is different from that of the effective schools researcher, seeking statistical outliers and then inferring the reasons for their success.

o The JDRP examines educational programs identified explicitly, and relies upon statistical comparisons to control groups specified by design. The effective schools researcher examines entire schools, perhaps best conceived as complex constellations of programs, policies and practices, and relies upon statistical comparisons to a population of similar schools.

o The two kinds of problems bring different types of statistical evidence to bear, and in different ways. Those schools or programs judged effective by the JDRP criteria and those identified by school effectiveness criteria would probably have little overlap. Differing expectations in focus and scope would dictate this divergence. Every single program in an effective school might fail the JDRP criteria. In another school, several programs might be judged effective through the JDRP criteria, and yet the school might well fail the test of effectiveness.

It is worth noting that a number of early Title III projects were total school innovation efforts, not unlike effective schools projects. As might be expected, these projects are not to be found in NDN.

As the NSTA/NSF and effective schools examples indicate, there are other ways in which the recognition task can be conceptualized, with the results depending on the approach taken. Each favors certain kinds of projects at the expense of others; given their different emphases, this is not surprising. The critical issue is to first understand the nature of the project to be evaluated, and then to determine the most meaningful way of recognizing quality for that type of project, and for the type of decision to be made.

Matching the JDRP Process to Educational Standards

A second means of analyzing the JDRP is through the use of the Standards for Evaluations of Educational Programs, Projects and Materials, developed by the Joint Committee on Standards for Educational Evaluation. Of course, the JDRP was not designed with the Standards in mind, but since the Standards represent a consensus of what good evaluation practice should be, it is reasonable to expect a substantial agreement. The thirty standards prepared and tested by the Joint Committee are grouped in four general areas: Usability, Feasibility, Propriety, and Accuracy. These cover a full range of concerns that might be raised about any program evaluation. Description and examples are provided with each Standard to aid in applying it to an individual evaluation.

The Standards provide an interesting reference point for the JDRP, because the JDRP primarily reviews a project's evaluation and then makes a judgment from this and the results reported by the evaluation. Although the ultimate decision is to approve or disapprove the project, the JDRP does not review the substance of the project, and usually has only a limited knowledge of its nature.

The JDRP process is perhaps best seen as a meta-evaluation that places certain expectations for information from projects through its Guidelines, and then reviews the adequacy and compellingness of that documentation and information provided in the oral presentation. It is also important to note that although the JDRP sets program evaluation standards for projects, the JDRP does not conduct a program evaluation. The meta-evaluation is more of a quasi-legal process for decision-making that operates under a totally different set of procedures than a project.

What we have done is to apply the Standards to both the expectations of JDRP for project evaluation submissions, as reflected in their Guidelines, and the JDRP process itself. The latter could be done only to the extent that we could interpret the process in light of program evaluation standards. Results of this analysis are reported in Table 1. In order to keep the analysis as simple as possible, for each Standard we simply made a judgment of YES the Standard was being employed, or NO, it was not. In doing this, we took into consideration the purpose of the JDRP and gave it credit for reflecting the Standards even if the Standards are stated in more general terms than the JDRP mandate.

For example, several of the Standards speak to the need for being responsive to multiple audiences and multiple purposes. It was our judgment that the JDRP mandate required a more narrow focus by the project. However, where the mandate did not require the approach adopoted, we applied the Standard as directly as possible.

In Column I of Table I, the JDRP's expectations of projects are summarized.. The JDRP matches well with the Standards in two categories: Usability and Accuracy. For Usability, only Evaluator Credibility and Report Dissemination were given NOs(the Guidelines do not address either of these) while all eleven standards in the Accuracy category are addressed. There is considerably less agreement in the other two categories: Feasibility and Propriety. None of the Feasibility Standards and only two of the Propriety standards-Full and Frank Disclosure and Rights of Human Subjects- are addressed, the latter only indirectly through the program warranty.

Thus it is clear the JDRP has selected to review only certain kinds of standards while ignoring others. Further, this selection is not part of the mandate. Feasibility and Propriety could have been included, but probably were not because they are less outcome-oriented than Accuracy and Usability. Lack of emphasis on these other concerns, particularly Feasibility, has led the JDRP to adopt an abstract view of evaluation, as if projects could set out to create an ideal evaluation without constraints. It has also spawned a cottage industry of evaluators brought in by projects to write submissions to make it appear this was the case.

Looking next at the JDRP process itself,as summarized in Column II, we find a strikingly different pattern. Overall, Usability is at the same high level, but in both the Feasibility and Propriety categories there is also a close match with the Standards. For Accuracy there is considerably less match, which can be explained in several ways. First, the Accuracy standards proved the most difficult to apply to the JDRP process, because many are stated in terms that are appropriate only to program evaluation. For this reason, we chose not to apply four of the Accuracy standards to the JDRP process. Another reason for the lower emphasis on Accuracy in the JDRP process is the limitations that exist in its review. There is no reason to believe that the JDRP could aspire to the same level of accuracy through several hours of reading and perhaps an hour of oral discussion that projects obtain through several years and thousands of dollars of evaluation activity. However, this is not intended to dismiss any concerns regarding the accuracy of the JDRP process. It is our belief that there are problems with accuracy, the most glaring being the absence of evidence regarding reliability and validity of the decisions made.

Table I
Comparison of Standards to the JDRP Expectations
for Projects and the JDRP Process of Review

Standards*	I Project Submission	II JDRP Process
Usability		
A1 Audience Identification	YES	YES
A2 Evaluator Credibility	NO	YES
A3 Information Scope and Selection	YES	YES
A4 Valuational Interpretation	YES	NO
A5 Report Clarity	YES	NO
A6 Report Dissemination	NO	YES
A7 Report Timeliness	YES	YES
A8 Evaluation Impact	YES	YES
FEASIBILITY		
B1 Practical Procedures	NO	YES
B2 Political Viability	NO	YES
B3 Cost Effectivness	NO	YES
PROPRIETY		
C1 Formal Obligation	NO	YES
C2 Conflict of Interest	NO	YES
C3 Full and Frank Disclosure	YES	NO
C4 Public's Right to Know	NO	YES
C5 Rights of Human Subjects	YES	YES
C6 Human Interactions	NO	YES
C7 Balanced Reporting	NO	YES
C8 Fiscal Responsibility	NO	YES
ACCURACY		
D1 Object Identification	YES	NO
D2 Context Analysis	YES	NO
D3 Described Purposes and Procedures	YES	YES
D4 Defensible Information Sources	YES	**
D5 Valid Measurement	YES	NO
D6 Reliable Measurement	YES	NO
D7 Systematic Data Control	YES	**
D8 Analysis of Quantitative Information	YES	**
D9 Analysis of Qualitative Information	YES	**
D10 Justified Conclusions	YES	YES
D11 Objective Reporting	YES	YES

* Joint Committeee on Standards for Educational Evaluation, Standards for Evaluations of Educational Programs, Projects and Materials. New York: McGraw Hill, 1981.
** We were unable to apply this Standard, as written to the JDRP

Conclusions

1. The JDRP has been an important part of educational evaluation for more than a decade, both because of the number of projects it has reviewed, and the gate-keeping function it plays for the National Diffusion Network.

2. In fulfilling this role, the JDRP has adopted a particular point of view and emphasis in evaluation- the judgment of impact as determined through the quasi-experimental methods of the social sciences. Recognition systems that have adopted other criteria and methods have achieved different results. While the JDRP process may have served well in judging the quality of many projects, it would not have served as well in judging science or effective schools projects.

3. Analysis of the JDRP through the educational evaluation Standards highlights its strong emphasis on Accuracy, and the lack of emphasis in the areas of Feasibility and Propriety. Lack of interest in Feasibility has been particularly difficult for the practicing evaluators to accept, as it has made it impossible for them to address the constraints under which they have had to work, or to argue that the evaluation conducted may have been the best possible under the circumstances.

4. In comparison to its expectations of the projects it reviews, the JDRP process fits a notably different set of standards. The JDRP process places more emphasis on the practical considerations of Feasibility and Propriety, but does not meet the Accuracy standards expected of its projects. Of particular concern is the almost total absence of information regarding the reliability and validity of its decisions. Some of these studies would not have been difficult to do, and should have been carried out.

5. Because of its emphasis on outcomes determined in particular ways, the JDRP has locked itself into a strategy that emphasizes only a certain kind of project. As the numbers of this type of project have diminished, due particularly to the demise of Title III, JDRP has allowed itself to become obsolete. Its major role is now to recertify projects it has previously approved. Any learning about the nature and mechanics of recognition that the JDRP might have brought to new situations has almost certainly been lost.

6. While the JDRP has been excellent in providing timely feedback to the projects coming before it, and to the NDN, it has failed to provide much information about itself or its deliberations to a broader audience in education and evaluation. If we are correct that the JDRP has been a significant element in the practice of program evaluation, it has also been one of the best kept secrets.

Because individual project deliberations were not documented in some detail, the JDRP has treated each project as one-of-a-kind, and the thought that went into its deliberations has been lost. There has been no attempt to build a rationale for why decisions were made similarly or differently because there was no reference to earlier decisions. While employing a quasi-legal model, the JDRP failed to build any case law or precedents.

The lack of a case law of decision rationales has had several implications. First, the JDRP has lost the opportunity for the internal monitoring this would have provided. Precedents place constraints on what decisions can be made,requiring strong justifications for actions against precedent. Second, the field of evaluation has lost the opportunity to learn about the types of distinctions and conditions that the JDRP weighed. JDRP does not deal with ideal examples. It deals with real-life examples that fall short of the ideal. It is the nature of the shortcomings that were either approved or not approved, and the reasons for those actions, that would have been helpful to practicing evaluators. It is a fair interpretation that the lack of any precedents on which to base their submissions is one of the greatest complaint evaluators have with the JDRP.